

The Effect of Ambient Scent on Consumers' Perception, Emotions and Behaviour – a Critical Review

Christoph Teller[†], Charles Dennis^{††}

[†]...[corresponding author] Surrey Business School, University of Surrey, Guildford, GU2 7XH, UK, email: c.teller@surrey.ac.uk, Phone: +44 (0) 1483 68 3981.

^{††}... Lincoln Business School, University of Lincoln, Brayford Pool, Lincoln, LN6 7TS, UK, email cdennis@lincoln.ac.uk, Phone: +44 (0)1522 835686, Skype charles.dennis1, Mobile/cellphone +44 (0)7984 92113

Abstract

The effect of olfactory stimuli on consumer behaviour has received little attention in marketing and retailing literature compared to other atmospheric cues. Researchers report ambiguous findings and shortcomings of measurement approaches. Based on a critical literature review, a field experiment in a regional shopping mall investigates the effectiveness of odour. Before-and-after surveys of randomly-selected shoppers in experimental and control groups were conducted and different experimental designs simulated. Those designs not controlling either extraneous errors or attitudinal differences between control and experimental group reveal a positive effect on factors operationalising mall perception and/or consumers' emotions. The design controlling both sources of bias indicates no impact of odour on the dependent variables. None of the behavioural variables were affected in any case. This paper questions prior findings regarding the effectiveness of odour in a shopping mall environment and calls for more rigour in investigating the effectiveness of atmospheric stimuli in general.

Keywords

Odour, Aroma, Atmosphere, Shopping Mall, Experiment, Consumer Behaviour

Biographies

Christoph Teller is a senior lecturer in marketing in the Institute for Retail Studies, which is in the Stirling Management School at the University of Stirling. He was previously assistant professor at the Vienna University of Economics and Business. Christoph's research interests are in the area of store (format) and agglomeration (format) choice/patronage, retail logistics and web-based research methodologies.

Charles Dennis is a Reader at Brunel University, London, UK. His teaching and research area is (e-)retail and consumer behaviour. Charles is a Chartered Marketer and Fellow of the Chartered Institute of Marketing. Charles's publications include *Marketing the e-Business*, (1st & 2nd editions) (joint-authored with Dr Lisa Harris); the research monograph (combined textbook) *e-Retailing* (Routledge) and *网上零售理论与实务* 中国物资出版社 (China Logistics Publishing House); and research monograph *Objects of Desire: Consumer Behaviour in Shopping Centre Choice* (Palgrave). His research into shopping styles has received extensive coverage in the popular media.

Introduction

The phenomenon 'atmosphere' and its impact on consumer behaviour has widely been considered in research over decades (Turley and Milliman 2000). Atmosphere can be regarded as a multi-faceted, latent construct which can be decomposed into different atmospheric stimuli or cues. These stimuli are proposed to be (1) recognisable by consumers' senses, (2) can lead to an intended reaction and (3) be effectively manipulated by retailers (Turley and Chebat 2002, Kotler 1973). Aroma is one of few marketer-manipulable atmospheric stimuli to have been studied yet amongst retail atmospherics studies, few have concerned aroma (Turley and Milliman 2000; Chebat and Michon 2003). Bone and Ellen (1999) critically discusses the effectiveness of aroma with respect to possible consumer reactions. This includes the affective and evaluative response to an object, (e.g. product, brand, store), behavioural intentions and actual behaviour.

Nonetheless, the body of literature dealing with the evaluation of these effects can be regarded as quite fragmented. Some articles focus on the effect regarding certain kind of products or brands (Bone and Jantrania 1992; Mitchell et al. 1995; Morrin and Ratneshwar 2000), some on products and stores (Spangenberg et al. 1996; Orth 2005), most of them regarding retail store or other kind of store-like environments (Mattila and Wirtz 2001; Gueguen and Petr 2006; Hirsch 1995; Knasko 1989) and an increasing number on shopping malls (Chebat and Michon 2003; Michon et al. 2005). The measurement approaches or experimental designs to evaluate the relationship between the use of aroma and the proposed dependent variables can also be seen as different. This includes the selected test units, the experimental environment, the sample selection procedure, the type of stimulus (aroma) etc. Furthermore, the results are ambiguous. A number of papers present insignificant, minor or only indirect effects of odour regarding attitudinal and in particular behavioural variables (eg. Gueguen and Petr 2006; Chebat and Michon 2003; Orth 2005; Mattila and Wirtz 2001).

Consequently, the question whether 'smell sells' or whether the use of olfactory stimuli results in reactions that are desirable from the point of view of marketing managers cannot yet be answered with a definite 'yes'. Some researchers provide quite logical justification for this unsatisfactory state of the art of literature. For example, Ward, Davies and Kooijman (2002); Bone and Ellen (1999); or Gulas and Bloch (1995) identify sets of variables that might moderate the effect directly and indirectly. These include several atmospheric cues including other olfactory stimuli, individual (e.g. demographic or psychographic) characteristics of consumers, congruency of scent, past experience, physiological predispositions, scent preference or perception of scent. Notwithstanding, only few studies include any of these moderators and confirm their relevance (e.g. Michon et al. 2005; Orth 2005; Chebat and Michon 2003).

The diversity of findings raises the question of whether the effectiveness of ambient scent can be taken for sure based on the state of the art of literature. The basic research question is consequently: What consumer reactions can be identified when (ambient) scent is used in a retail environment? Besides addressing the research question the aims of this paper are (1) to critically evaluate existing results and applied methodology in marketing and retail related research on the effectiveness of odour in a shopping environment, (2) empirically test the effect of odour by applying different experimental approaches and – based on that (3) discuss the current state of the art of olfaction research in marketing and retailing with respect to the identified shortcomings in literature. Thus the contribution of this research is to question both the validity of existing findings on and the appropriateness of applied experimental approaches to measure the effectiveness of odour – with an explicit focus on marketing and retail research.

The realm of the paper is as follows: based on this introduction we discuss the findings from literature regarding the effects of odour on shopping behaviour and the different research designs applied. A conceptual framework is developed and used to frame hypotheses and operationalise variables. Next, the research design and the results of an empirical study are

presented, then discussed with respect to prior findings. The paper concludes with a short summary and an outlook for further research.

Literature Review

In line with the research question, this literature review focuses mainly on those publications dealing with the effect of aroma in a retail related context. Turley and Milliman's (2000) synthesis of publications on atmospheric effects show that aroma had not received enough attention compared to other atmospheric cues. Only three out of sixty empirical studies had been identified that focused on aroma as an independent variable. Almost a decade later the situation is little changed. We identified just eight more journal publications in refereed journals that investigate the effect of aroma on consumer behaviour related reactions (see Table 1).

In all of these articles concerning aroma, (quasi) experiments had been conducted. This literature review evaluates the state of the art of olfactory marketing and retailing research with reference to the research designs.

Research Context

Whereas older publications (pre-2002) look at the effectiveness of aroma in a product and/or store context, more recent studies focus on the supra-store context of shopping malls (Michon et al. 2005; Chebat and Michon 2003). This might be because of the rising number of such retail (agglomeration) formats and preference of both consumers and retailers for these for these retail environments. Furthermore, atmosphere (as a set of stimuli) has been identified to be of determinant importance for retail agglomeration attractiveness (e.g., Teller 2008; Teller and Reutterer 2008; Teller et al. 2008). Nonetheless, the complexity of identifying and measuring relationships between an aroma and a consumer response increases due to the high number of extraneous variables such as other atmospheric stimuli and influences and/or the heterogeneity of different kinds of clientele patronising different stores (Teller and Reutterer 2008; Van Kenhove et al. 1999).

Experimental Location and Test Units

A considerable number of publications investigate the phenomenon in a product and brand context by conducting laboratory experiments and/or using students as their subjects (Orth 2005; Morrin and Ratneshwar 2000; Mitchel et al. 1995; Spangenberg et al. 1996; Bone and Jantrania 1992). The applied experimental approach offers a high control of the research environment, leading to a high internal validity and consequently is time and cost effective in particular when the participants are students (Malhotra and Birks 2007). On the other hand, such an artificial research environment may cause reactive errors or measurement artefacts and, furthermore, the external validity is limited because of the focus on a very specific customer group, i.e. students of a particular university (Malhotra and Birks 2007). In other words, the findings regarding the effectiveness of aroma can be hardly drawn to other populations and thus have a more exploratory character.

Sample Selection Procedure

Although most of the identified publications include actual customers as subjects, only three stated that their sample had been selected randomly (Michon et al. 2005; Orth 2005; Chebat and Michon 2003). Three of them observed the whole population over the defined research period (Gueguen and Petr 2006; Hirsch 1995; Knasko 1989). A convenience sampling approach was applied in the others and thus the results could be affected by sample selection bias and can therefore be generalised only for the investigated sample.

Experimental Design

All research designs investigate the effects of odour by observing and/or surveying two groups of test units. One group that had been exposed to the stimulus (i.e. experimental group) and one group that had not (i.e. control group). Thus, the design which applied the random sampling technique can be seen as true experimental design, denoted as the 'after

only with control group' whereas all the others are 'static group comparison'. Due to the lack of randomisation of the selection procedure this latter pre-experimental design fails to control the above mentioned 'extraneous' variables which are quite numerous in a store or even agglomeration context. Although the 'after only with control group' design does not – theoretically – suffer from this problem although it is sensitive with respect to sample mortality and selection bias (Malhotra and Birks 2007). Even more critical is the underlying assumption that the randomly selected groups are equal in any particular respect with respect to the pre-treatment measures (Malhotra and Birks 2007). Since there is no pre- or before-measurement this assumption is left unproven. In a retailing context – store or agglomerations – a considerable number of variables can act as moderators on the investigated effects if they are different between the control and experimental group. Socio-demographic, socio-economic, psychographic but also attitudinal variables can have such an effect, making controlling for these variables difficult (Bone and Ellen 1999; Gulas and Bloch 1995).

The only way to confirm this assumption of group homogeneity is to apply a 'before-after with control group design' which has not been used in any of the presented articles. Finally, it is worth mentioning that only a few publications consider selective moderators like retail density (Michon et al. 2005), pleasantness of scent (Orth 2005) or music (Mattila and Wirtz 2001).

Endogenous Variables

The aromas (exogenous variables) used in most studies were citrus (pleasing, arousing or stimulating) and lavender (neutral) (based on the findings of Spangenberg et al. 1996). The endogenous variables proposed to be directly or indirectly affected can be divided into three groups:

- Perceptual/attitudinal variables: overall perception or perception of certain attributes (i.e. quality) of products, stores or malls, store environment,
- Emotional variables: pleasure, arousal, stimulation level
- Behavioural and intentional variables: actual and perceived spending and retention time, intended spending and retention time, variety-seeking and curiosity-motivated behaviour, information search, choice behaviour.

The inclusion of the above stated variables depended on the (implicit) research question of each study and also on the theoretical framework applied. Additionally, the operationalisation of variables varied across the different studies. Therefore, it is difficult to compare and confirm the effects identified as significant or insignificant (see Table 1).

Referring back to the previous section, the measurement of changes in the endogenous variables caused by the exogenous variables can be problematic if there is no pre-measurement. For example changes in the emotional state, e.g. pleasure or arousal, can be different for each test (person) unit before they are exposed to the experimental treatment. The same is true for the attitude and historic perception of attributes of an investigated object, e.g. store. Furthermore, the individual shopping situation and task or personal characteristics like disposable income can predetermine the shopping behaviour in the experimental environment (see e.g., Van Kenhove et al. 1999; Kahn and Schmittlein 1992). Thus, a causal interpretation of differences between control and experimental group with respect to the stimulus might be affected by such preliminary differences between test units. Since in none of the studies a pre-measurement took place, the causalities of investigated effects are to some degree questionable.

[Take in Table 1 about here]

By summing up the discussion of the identified literature and the applied experimental approaches it can be said that the more recent publications apply more rigorous approaches by utilising random sampling (Michon et al. 2005; Orth 2005; Chebat and Michon 2003). Even so, the problem of measuring the effectiveness of (ambient) scent in a retail environment calls for more complex and consequently laborious experimental designs which incorporate a pre-measurement phase. This is because of the high number of extraneous

variables affecting test units (people) and the heterogeneity of test units themselves in terms of demographic, psychographic, attitudinal and behavioural variables. In summary, on the basis of prior research, the effectiveness of (ambient) scent cannot be guaranteed. This supports the call from Evanschitzky et al. (2007) and Neuliep (1991) regarding the demand for critical replication studies. In that respect, Evanschitzky et al. (2007) cautioned practitioners that 'scientific findings rest upon replication ... few results in marketing have been successfully replicated... given these results, practitioners should be sceptical about making decisions based on the findings of the predominantly single-shot studies reported in the leading marketing journals' (2007, 413).

Conceptual Framework

As a basis for a reference study we developed a conceptual framework within which we set up three hypotheses that represent the most frequently investigated effects of odour. According to the aims of this paper the hypotheses are consequently tested by applying the whole spectrum of experimental designs as this has been done in previous studies.

In the formulation of the hypotheses we focus on a test environment that faces growing preference from consumers' and consequently retailers' points of view – the shopping mall (Author 2010). Recent research demonstrates the extraordinary importance of the perceived atmosphere – including scent - on shopping mall attractiveness (Teller 2008). Furthermore, two of the most rigorous studies conducted in terms of measuring the effectiveness of ambient scent focus on the same research environment. Consequently, the papers of Michon et al. (2005) and Chebat and Michon (2003) serve as templates for both the hypotheses and the empirical research design.

In line with most of the publications on the effectiveness of odour in a shopping environment (see e.g., Spangenberg et al., 1996) we apply the Stimulus-Organism-Response model of Mehrabian and Russel (1974). Therewith, we include two constructs to measure the organism, i.e. perception of the object and emotions of the test units, and one the response, i.e. customer behaviour (McGoldrick and Pieros 1998). Within this conceptual framework we only focus on direct effects between aroma and the proposed dependent variables since their existence is crucial for the existence of indirect or mediating effect of emotions or perceptions on shopping behaviour. The most commonly investigated effects can be described as follows.

Bitner (1992) proposes effects of environmental cues, e.g., in terms of atmospherics, on consumers' perceptions and thus evaluation of objects, having an influence of consumers' cognition. More specifically, Spangenberg et al. (1996) proposes and provides empirical proof for odour for having an impact on the perception of products and a store. Furthermore the influence of odour on the perception of a product, a store or a mall – was investigated by the following studies presented by Michon et al. (2005), Chebat and Michon (2003), Mattila and Wirtz (2001), and Bone and Jantrania (1992). There the effect is measured by the evaluation of object attributes. In terms of a shopping mall these attributes comprise e.g. the retail tenant mix, the non-retail tenant mix, price-value ratio and product range of merchandise, personnel and overall atmosphere (Teller, 2008). The perception of the mall and the evaluation of mall attributes can be seen as a core-determinant of the mall related consumer behaviour (Finn and Louviere, 1996). We therefore propose the following:

H₁ [Stimulus → Organism]: Aroma has a significant, positive impact on consumers' perception of a shopping mall.

Donovan et al. (1994) propose effects between environmental stimuli and emotional states. Within this framework of the modified Mehrabian-Russel environmental psychology model the effects of aroma on consumers' emotions were investigated in the studies of Michon et al. (2005), Chebat and Michon (2003) and Morrin and Ratneshwar (2000). There emotions are seen to have a mediating role and consequently affect the perception of an object or shopping behaviour on site (Donovan et al. 1994). Most frequently, the latent construct emotions have been operationalised by pleasure and arousal measured by the scale of Mehrabian and Russel (1974). In accordance with Chebat and Michon (2003) we assume

that odour has a positive effect on both pleasure and arousal. In the context of a shopping mall we set up our second hypothesis:

H₂ [Stimulus→Organism]: Aroma has a significant, positive impact on consumers' emotions in a shopping mall.

Gulas and Bloch (1995) propose an effect between the use of odour and affective response of consumers in terms of their behaviour. The consumer behaviour as a dependent variable has been considered most frequently in empirical studies although a direct effect was not always proposed (Gueguen and Petr 2006; Michon et al. 2005; Chebat and Michon 2003; Mattila and Wirtz 2001; Spangenberg et al. 1996, Hirsch 1995, Knasko 1989). Indicators that are frequently used to operationalise the construct of shopping behaviour in a mall context are (planned or actual): spending in (non-)retail stores; retention time; and number of stores visited. Based on that we derive our last hypotheses:

H₃ [Stimulus→Response]: Aroma has a significant, positive impact on customers' behaviour in a shopping mall.

To test the hypotheses by using different experimental designs we conducted the following empirical study.

Empirical Study

Experimental Design

In order to investigate the impact of olfactory stimuli on shopping behaviour, we conducted a field experiment in a regional, centrally located shopping mall in a central European city. The particular mall was chosen because of its small size (30,000 m², 40 tenants, 680 parking spaces, 500 employees) and its design that can be considered as state of the art (ICSC European Shopping Centre Award Winner 2008).

The field experiment contained two steps where we applied a survey approach including observational elements. The applied 'in-vivo' survey approach has the advantage that the phenomenon is investigated in a biotic shopping situation (Teller and Reutterer 2008). It enables the researcher to confront respondents with questions about their actual perceptions, emotions and behaviour on the respective shopping trip. Despite the resulting high internal validity of the approach the control of all extraneous factors on the experiment, e.g. weather, mood of the respondents etc., is almost impossible. Nevertheless, we see the disadvantages of the applied approach outweighed by the advantages to investigate the effectiveness in a real life situation (Michon et al., 2005).

In the first step, customers who entered the mall during a period of one week were surveyed by using an interviewer-administered questionnaire. Thereby we used a time sampling approach and selected those customers who entered the mall every full quarter of an hour and passed an exactly defined point at the entrance area. The number of respondents varied according to the time of day (Sudman 1980). The test units were confronted with questions regarding their shopping behaviour in the mall in general, about their actual shopping in the mall (including task definition, planned spending, shops to be visited, time to be spent). Thereafter, they had to evaluate their emotional state and the mall based on 52 attributes according to Teller et al. (2008). After a self-characterisation based on the demographic and socio-economic questions, respondents were asked to return after they had completed their shopping in that mall. At this stage of the interview, the time was recorded on the questionnaire to obtain objective measures regarding the retention time. Before leaving the mall, respondents reflected on their shopping trip in the mall based on questions regarding perceived retention time, actual stores visited and emotional state. Finally, they were once more confronted with the same scale of attributes characterising the mall.

The second step of the experiment included the same sampling and interview procedure using the same research instrument. In that second week an aroma was spread in the mall. This was arranged by a professional aroma marketing agency with wide experience in that field. The terminals were located in the common area of the mall and taking into account the size and air-circulation. No aroma was used by all the other 40 tenants and interviewers were

(again) told not to wear any perfume (Chebat and Michon 2003). The aroma consisted of a mixture of orange, grapefruit, bergamot, cinnamon, cardamom, ginger, pimento and other additives. The characteristics of that aroma are widely used in comparable retail settings and other studies and are described as warming, stimulating, sweet and citric-like. It is finally worth mentioning that the weather and thus the light intensity did not vary over the two weeks in June. Overall it can be said that there were no major extraneous variable influencing the comparability of the two weeks in terms of the use and effectiveness of the used odour.

One hundred and thirty-six usable questionnaires were obtained from the 'control group' where no aroma was used and 176 from the 'experimental group' which was exposed to the aroma. The sample mortality included 38 respondents who were reluctant to answer the questions after their visit. This group turns out not to be significantly different to the final sample with respect to the data collected in the pre-measurement. At no stage of the interviews were the respondents told about the aim of the study, i.e. test the effectiveness of aroma.

Characterisation of Respondent Groups

The two respondent groups can be regarded as homogenous with respect to their demographics and shopping behaviour in the mall since no significant differences could be identified (see Table 2). Due to the 'in vivo' survey approach and the sampling procedure, the selected groups of customers do not totally reflect the demographic structure of the whole population of the urban area. The surveyed clientele of the mall can be characterised by being dominated by women and highly educated. The sample contains a remarkably large group of students and senior citizens (see also the standard deviation of the age figure) which can be explained by the demographic structure of the specific district in which the mall is located. Our respondents were experienced shoppers with respect to the mall since, on average, they spend more than 6 visits per month there, stay longer than three quarters of an hour per visit and shop at more than two outlets there.

[Take in Table 2 about here]

Analysis

Simulated experimental designs: Despite the theoretical and empirical foundations of the three hypotheses methodological problems can be identified in some prior studies. Thus the accuracy of the supporting results is to be questioned. The shortcomings account for (1) the lack of controlling extraneous variables which are numerous in such a complex retail environment like a store or a shopping mall and (2) the lack of testing the homogeneity-assumption regarding the two groups of test units (i.e. experimental and control group).

The data derived from our empirical study enables us to test the hypotheses as if by different experimental designs (Figure 1). The most basic one can be denoted as 'one group pretest-posttest' design. Although this test has not been applied in any of the identified prior studies varying results compared to the other two designs reveal the impact of extraneous variables. The effect of the stimulus is investigated by looking at differences between the pre- and post-measurement of the dependent variables ($\Delta 1$). The 'Posttest only control group design' investigates the differences between the control and the experimental group regarding the proposed dependent variables ($\Delta 2$). This would correspond to the 'static group design' if no random sampling procedure was adapted. The last and most extensive approach is the 'pretest-posttest control group' where the differences between the control group ($\Delta 3_a$) and the experimental group ($\Delta 3_b$) are compared ($\Delta 3$). There, the group-homogeneity can be tested by looking at the differences regarding the pre-measurement ($\Delta 3_0$).

[Take in Figure 1 about here]

Tests of significance: Mean values were calculated of all items standing behind the perceptual and emotional (latent) variables. This procedure can be justified by the satisfactory internal consistency of each variable (Cronbach's α around or above 0.70; see

appendix). In order to identify an effect of the olfactory stimuli, rank-sum tests were conducted between the variable or values of the respective groups. For independent samples, e.g. comparison between the experimental and the control group, we applied the Mann-Whitney U-Test. The Wilcoxon-Test was used for dependent samples, e.g. comparison of the before and after evaluation of the experimental group. The use of these particular statistical tests was motivated by the ordinal or 'only' quasi-metric measurement level of the applied rating scales and fewer assumptions regarding the distribution assumption of our data – in particular with respect to the behavioural variables (Field, 2009).

Results

One Group Pretest-posttest Design ($\Delta 1$)

First, only the experimental group is taken into account and the perception of mall attributes, respondents' emotions and shopping behaviour is compared between the pre- and post-measurement (see Table 3). It can be seen that the ratings of the 'retail tenant mix', the 'non-retail tenant mix', the 'price-value ratio' and the 'personnel' of the mall are rated significantly higher after the visit to the mall than before (Wilcoxon-Test; $p < .05$). Pleasure and arousal did not change during the visit. Interestingly, the shopping behaviour shows significant changes but not in favour of the mall tenants. The actual spending for food and drinks, the actual retention time and the actual number of stores visited was significantly lower than had been planned when entering the mall. Consequently, only H_1 can be confirmed with respect to the significant variables.

[Take in Table 3 about here]

Again, it needs to be mentioned that although significant differences are investigated, these results cannot necessarily be interpreted as demonstrating that the aroma caused these changes during the mall visit – as we will demonstrate with the 'pretest-posttest control group design' in the section, below.

'Posttest Only Control Group Design' ($\Delta 2$)

This design requires the inclusion of the results of the control group (see Table 4). Again we see that selected perceptual variables show significant differences (Mann-Whitney U-Test). The '(non-)retail tenant mix', the 'product range', 'personnel' and 'other atmospherics' were rated higher by the experimental group. Furthermore the test units of the scented group feel significantly more aroused. No differences regarding the behavioural variables can be identified except the number of stores visited, which is significantly higher for the control group.

[Take in Table 4 about here]

This design can be regarded as the most sophisticated one presented in prior literature and on that basis, H_1 and H_2 would be partially supported. Nevertheless, since no pre-measurement is included in investigating the 'causal' effect of the olfactory stimulus, the homogeneity of the control and experimental group can only be assumed but is left unproven. Therefore, we do not accept H_1 and H_2 as partially supported but rather, turn to the 'Pretest-posttest control group design' in Section 5.3, below.

'Pretest-posttest Control Group Design' ($\Delta 3$)

The final design incorporates both a pre-measurement and a control group. Before presenting the results from comparing the changes observed ($\Delta 3_a$ and $\Delta 3_b$) between the two groups, the homogeneity assumption is tested with respect to a broad set of characterising variables ($\Delta 3_0$). In other words we assure whether the random sampling approach really produced two homogenous samples – in any relevant respect.

By applying simple bivariate statistical methods we see that the groups are homogenous ($p > .1$) with respect to the following variables regarding the pre-measurement:

- Personal characteristics (see Table 2)
- Average shopping behaviour related to the mall under investigation (see Table 2)
- Shopping situation: Task definition (χ^2 -Test; according to Van Kenhoven et al. 1999), Involvement (Mann-Whitney *U*-Test, scale according to Wakefield and Baker 1998)
- Overall perception of mall attractiveness: Satisfaction, retention proneness and patronage intention (Mann-Whitney *U*-Test, scales according to Teller and Reutterer 2008)
- Planned shopping behaviour on site: spending, number of stores to be visited (Mann-Whitney *U*-Test)
- Perception of mall attributes: Non-retail tenant mix, price-value ratio, smell and other atmospheric stimuli (Mann-Whitney *U*-Test)
- State of emotions: pleasure (Mann-Whitney *U*-Test)

Nevertheless, we have to reject the group-homogeneity assumption because we face the following significant differences ($p > .1$):

- Planned shopping behaviour on site: retention time (higher ratings in the control group, Mann-Whitney *U*-Test)
- Perception of mall attributes: retail tenant mix, product range, personnel (higher ratings in the experimental group, Mann-Whitney *U*-Test)
- State of emotions: arousal (higher ratings in the experimental group, Mann-Whitney *U*-Test)

It can be concluded that the homogeneity assumption cannot be confirmed for five variables included in the conceptual model. Recall, interestingly, that all perceptual variables and the emotional variable turned out to be affected by aroma when tested with respect to the 'Posttest only control group design'. Consequently, those results are questionable.

Pretest-posttest control group design is not sensitive regarding this heterogeneity because the pre-measurement serves as a baseline or reference value for each test unit. The effect of aroma is then investigated by first calculating the differences between the post-measurement and the reference value for both groups and then comparing these values between the groups (Table 5).

[Take in Table 5 about here]

The results of the pretest-posttest control group design are sobering since none of the variables are significantly different in favour of the experimental group. The experimental test units actually visit less shops than they planned compared to their counterparts in the control group.

We also investigated whether positive (significant) results can be found from exemplar demographic segments such as gender, age or educational level. In general the findings were similar within the segments. In gender for instance the only significant differences for the males ($n=120$) was that aroma only had a positive impact on non-retail tenant mix ($p < .01$) and retention time ($p < .05$). Whereas for the females ($n=192$) there was a negative effect regarding number of stores visited ($p < .05$). For the age 25 and younger segment ($n=85$) the aroma had a positive influence with respect to the perception of the odour only ($p < .05$), whereas the 50 and older ($n=108$) the effect was negative for the number of stores visited. For educational level the effect of aroma was negative for higher educational level ($n=190$) and for the number of stores visited ($p < .05$).

The varying results derived from the application of the three experimental designs leads us to conclude that the experimental design impacts the results derived from testing the other hypotheses. There were few differences regarding the findings for the demographic segments as notwithstanding the occasional positives, overall there was little positive influence of aroma for any of the demographic segments.

The most interesting segmentation findings come from the post-hoc segments considering aroma not very important (rating between 0 and 6 ($n=109$) vs. very important (rating between 7 and 9 ($n=203$)). As might be expected there were negative findings on the effect of aroma on spending and number of stores visited ($p < .05$) for those respondents for whom aroma

was considered of not being important. Even for those for whom aroma was considered of being important the only significant positive effect was on the perception of the non-retail tenant mix.

Conclusions

Taking into consideration the specific characteristics of the research object, i.e. shopping mall, the results confirm some of the findings from literature depending on the simulated experimental design. An overview of the hypotheses testing can be viewed in Table 6.

[Take in Table 6 about here]

The most explicit finding is that aroma did not affect any observed or surveyed variable of consumer behaviour – no matter which experimental design was simulated. This clearly contradicts the findings from Gueguen and Petr (2006), Mattila and Wirtz (2001), Spangenberg et al. (1996), Hirsch (1995) or Knasko (1989). Since no effect was identified for our ‘Before-after with control group design’, a depth of rigour that has not been applied previously in this context, prior research that proposes indirect effects of aroma on behaviour also need to be treated with caution. Nonetheless, previous findings can be confirmed for the ‘post-test only with control group design’ (e.g. Michon et al. 2005; and Chebat and Michon 2003) and also where no pre-measurement had been included (e.g. Spangenberg et al. 1996; and Bone and Jantrania 1992).

[Take in Table 6 about here]

At the first sight, these results seem to be ambiguous but ultimately demonstrate how determinant the applied measurement approach is for the effects identified. The differences between the pre-experimental design and the most sophisticated design clearly show the strong effect from the numerous extraneous variables which in particular include other atmospheric stimuli in a mall. Consequently, the measurement of atmospheric stimuli without including a control group may result in false apparent support for hypotheses. Regarding the ‘post-measurement only with control group’ we have identified an even more substantial issue. The random sampling approach seemed at the outset to produce two homogenous sub-samples, i.e. the control and the experimental group. They are invariant regarding a number of variables like demographics, average shopping behaviour and the variables operationalizing the shopping situation. The two groups show significant differences regarding the dependent variables which might appear to be related to significant differences in five perceptual variables and one emotional variable. Nevertheless, by comparing with the result from the ‘pre-posttest with control group design’ it is clear that we are again confronted with false support for the hypotheses. As an aside we note that some prior work reports positive findings for particular demographic segments. Indeed these findings also indicate differences but when the most rigorous design is used our interpretation is that these differences are insufficient to conclude a positive effect of aroma. As a result, it can be concluded that the application of different experimental designs substantially affects the results.

Since none of the studies identified in academic literature incorporates a pre-measurement in identifying the effectiveness of the aroma used, we may question the conclusions that they report. This is especially true for those studies where a static group comparison had been used that report results regarding both the effectiveness and the ineffectiveness of aroma. I.e., results showing no effect of aroma on the dependent variables might also be different if a baseline measurement had been included. Indeed, perhaps as would be expected we find a difference with respect to the impact of aroma between the respondents considering aroma as important in contrast to those considering aroma as unimportant, viz for a full one third of the sample aroma had an negative impact on consumer behaviour.

Finally, it is worth mentioning that we do not conclude that the findings – from studies where ‘post-test only with control group’ was applied – are wrong but rather that they need to be

treated with care since the group homogeneity assumption did not appear to have been demonstrated. By referring back to our research question we tend towards a conclusion that aroma has no impact on either perception, emotions or behaviour of consumers. Nevertheless, the following limitations need to be taken into account when interpreting our finding.

Limitations and Outlook

Due to the specific character of retail settings regarding different geographical areas and times of the year, the findings suffer from limited external validity. Although a random sampling approach was used, the results can only be generalised to the clientele of the investigated regional mall over the research period.

In accordance with the purpose of our study and the sample sizes, we did not distinguish between certain customer segments. Therefore, it is possible that specific groups identifiable by demographic (e.g., gender and age), psychographic (e.g., hedonic or utilitarian shopping orientation) or behavioural characteristics (e.g., frequent or infrequent shoppers) might show different reactions to the aroma (Bitner, 1992).

The particular aroma used for this study can also be seen as a limitation. According to Gulas and Bloch (1995), scent preference or experience can moderate the effects investigated. That said, we were reluctant to use a different aroma to that used in most previous aroma studies.

Obviously, a shopping mall includes comparably more atmospheric stimuli compared to a store and thus the effective use of aroma turns out to be a quite complex task. The findings might well be different when the research object is a store or a single product. We justify our focus on the mall by the fact that the most rigorous prior studies have been conducted in a mall context and thus were most appropriate to be replicated.

In summary, this study has explored the replicability of prior studies of the effects of aroma in the shopping mall context and concluded that challenges to the findings of positive effects remain. In this instance, our findings do not support previous work on the positive effect of aroma although other contexts and other aromas could be investigated in the future. We recommend directing resources to more rigorous and extensive investigation than has been applied in the past not just for aroma but also in the search for other, perhaps more effective stimuli. Such stimuli might include music, visuals such as colours and in particular anything animated such as digital signage video screens.

References

- Bitner, M. J. (1992), Servicescapes: The Impact of Physical Surroundings on Customers and Employees. *Journal of Marketing*, Vol. 56, No 2, pp. 57-71.
- Bone, P. F. and Ellen, P. S. (1999), "Scents in The Marketplace: Explaining a Fraction of Olfaction", *Journal of Retailing*, Vol. 75, No. 2, p. 243-262.
- Bone, P. F. and Jantraia, S. (1992), "Olfaction as a Cue for Product Quality", *Marketing Letters*, Vol. 3, No. 3, pp. 289-296.
- Chebat, J. C. and Michon, R. (2003), "Impact of Ambient Odors on Mall Shoppers' Emotions, Cognition, and Spending: A test of competitive causal theories", *Journal of Business Research*, Vol. 56, No. 7, pp. 529-539.
- Evanschitzky, H., Baumgaarth, C., Hubbard, R. and Armstrong, J. S. (2007), "Replication Research's Disturbing Trend" *Journal of Business Research*, Vol. 60, No. 4, pp. 411-415.
- Field, A. 2009, *Discovering Statistics Using SPSS*. Los Angeles Sage.
- Finn, A. and Louviere, J. J. (1996), "Shopping Center Image, Consideration, and Choice: Anchor Store Contribution", *Journal of Business Research*, Vol. 35, No. 3, pp. 241-251.
- Gulas, C. S. and Bloch, P. H. (1995), "Right Under Our Noses: Ambient Scent and Consumer Responses", *Journal of Business and Psychology*, Vol. 10, No. 1, pp. 87-98.
- Guquguen, N. and Petr, C. (2006), "Odors and Consumer Behavior in a Restaurant", *International Journal of Hospitality Management*, Vol. 25, No. 2, pp. 335-339.
- Hirsch, A. R. (1995), "Effects of Ambient Odors on Slot-Machine Usage in a Las Vegas Casino", *Psychology & Marketing*, Vol. 12, No. 7, pp. 585-594.

- Knasko, S. C. (1995), "Pleasant Odors and Congruency: Effects on Approach Behavior", *Chemical Senses*, Vol. 20, No. 5, 479-87.
- Kotler, P. (1973), "Atmospherics as a Marketing Tool", *Journal of Retailing*, Vol. 49, No. 4, pp. 48-64.
- Malhotra, N. K. and Birks, D. F. (2007), *Marketing Research. An Applied Orientation*, Pearson Prentice Hall.
- Mattila, A. S. and Wirtz, J. (2001), "Congruency of Scent and Music as a Driver of In-store Evaluations and Behavior", *Journal of Retailing*, Vol. 77, No. 2, pp. 273-289.
- McGoldrick P. J. and Pieros, C. P. (1998), "Atmospheres, Pleasure And Arousal: The Influence Of Response Moderators", *Journal of Marketing Management*, Vol. 14, No. 1/3, pp. 173-197.
- Mehrabian, A. and Russel, J. A. (1974), *An Approach to Environmental Psychology*, MIT-Press Cambridge (MA).
- Mitchell, D. J., Kahn, B. E. and Knasko, S. C. (1995), "There's Something in the Air: Effects of Congruent or Incongruent Ambient Odor on Consumer Decision Making", *Journal of Consumer Research*, Vol. 22, No. 2, pp. 229-238.
- Morrin, M. (2000), "The Effect of Retail Store Environment on Retailer Performance", *Journal of Business Research*, Vol. 49, No. 2, pp. 167-81.
- Neuliep, W. (1991), *Replication Research in the Social Sciences*, Sage.
- Orth, U. R. and Bourrain, A. (2005), "Optimum Stimulation Level Theory and the Differential Impact of Olfactory Stimuli on Consumer Exploratory Tendencies", *Advances in Consumer Research*, Vol. 32, No. 1, pp. 613-619.
- Spangenberg, E. A., Crowley, A. E. and Henderson, P. W. (1996), "Improving the Store Environment: Do Olfactory Cues Affect Evaluations and Behaviors?", *Journal of Marketing*, Vol. 60, No. 2, pp. 67-80.
- Teller, C. (2008), "Shopping Streets Versus Shopping Malls - Determinants of Agglomeration Format Attractiveness from the Consumers' Point of View", *International Review of Retail, Distribution & Consumer Research*, Vol. 18, No. 4, pp. 381-403.
- Teller, C. and Reutterer, T. (2008), "The Evolving Concept of Retail Attractiveness: What Makes Retail Agglomerations Attractive When Customers Shop at Them?", *Journal of Retailing and Consumer Services*, Vol. 15, No. 3, pp. 127-143.
- Teller, C., Reutterer, T. and Schnedlitz, P. (2008), "Hedonic and Utilitarian Shopper Types in Evolved and Created Retail Agglomerations", *International Review of Retail, Distribution & Consumer Research*, Vol. 18, No. 3, pp. 283-309.
- Turley L. W. and Chebat J. C. (2002), "Linking Retail Strategy, Atmospheric Design and Shopping Behaviour", *Journal of Marketing Management*, Vol. 18 No. 1/2, pp. 125-144.
- Turley, L. W. and Milliman, R. E. (2000), "Atmospheric Effects on Shopping Behavior: a Review of the Experimental Evidence", *Journal of Business Research*, Vol. 49, No. 2, pp. 193-211.
- Van Kenhove, P., de Wulf, K. and van Waterschoot, W. (1999), "The Impact of Task Definition on Store-Attribute Saliences and Store Choice", *Journal of Retailing*, Vol. 75, No. 1, pp. 125-137.
- Ward, P., Davies, B. J. and Kooijman, D. (2003), "Ambient Smell and the Retail Environment: Relating Olfaction Research to Consumer Behaviour", *Journal of Business and Management*, Vol. 9, No. 3, pp. 289-302.

Source	Independent variable	Dependent variable [Mediators]	Experimental design	Probands and Sample size (n)	Research context	Significant (direct) positive effect
Gueguen and Petr (2006)	2 aroma	Retention time, spending	Field experiment, (observation)	All actual shoppers over a period of two weeks (n=88)	Restaurant	Retention time, spending, (only for lavender)
Michon et al. (2005)	Aroma	Mall perception, pleasure product quality perception, [retail density]	Field experiment, After only with control group design (survey)	Actual shoppers (n=9x31=279)	Community mall	Mall perception (medium), only at a medium density level
Orth (2005)	Pleasant and unpleasant aroma	Actual/optimum stimulation level, risk taking, variety seeking, curiosity-motivated behaviour	Laboratory experiment, after only with control group design (survey)	Persons from a consumer panel (n=248)	Store and product	Actual stimulation, only for pleasant aroma, risk taking and variety seeking, only for unpleasant aroma, curiosity-motivated behaviour
Chebat and Michon (2003)	Aroma	Pleasure/arousal, mall perception, product quality, spending	Field experiment, after only with control group design (survey)	Actual shoppers (experimental group: n=145; control group: n=447)	Community mall	Mall perception (low), product quality (low)
Mattila and Wirtz (2001)	Aroma, music	Pleasure/arousal, approach/avoidance, store environment, satisfaction	Field experiment, static group comparison (survey and observation)	Actual shoppers (n=270)	Store	Perception of store environment, impulse buying, satisfaction, only when music and aroma are congruent
Morrin and Ratneshwar (2000)	Aroma	Pleasure/arousal/dominance, brand perception, attention, memory	Laboratory experiment, static group comparison (survey and observation)	Students (n=50)	Brands	Brand evaluation time, recall (for unfamiliar brands)
Mitchell et al. (1995)	Aroma, product	Memory, information search, variety seeking behaviour, product choice	Laboratory experiment, static group comparison (survey)	Experiment 1: Students; n=77; Experiment 2: no characterisation, n=78	Product	Attention, variety seeking, information search, choice process, when aroma is congruent with product class
Spangenberg et al. (1996)	2 aroma	Evaluation of store, store environment, merchandise, specific products, patronage intention, buying intention, actual vs. perceived time spent, number of products examined	Laboratory experiment, static group comparison, (survey and observation)	Convenience sample including mostly students (n=704)	Store and product	Perception of store attributes, perception of store environment, perception of merchandise, purchase intention, retention time
Hirsch (1995)	2 aroma	Spending	Field experiment, (observation)	Actual gamblers at 18 slot machines (over three weekends (Saturday, Sunday)	Casino	Spending, only for one odorant
Bone and Jantrania (1992)	2 aroma	Overall product evaluation	Laboratory experiment, static group comparison, (survey)	Students (53)	Product	Overall product evaluation
Knasko (1989)	2 aroma	Evaluation of product attributes Spending, interaction with sales clerks, touching frequency of displays, retention time, [gender]	Field experiment, (observation)	All actual shoppers over a period of two weeks	Store	Retention time, spending, only for fruity/floral aroma

Table 1: Prior studies of the effects of aroma

Table 2: Respondents' profile

	Respondent groups		Control group (n=136)		Experimental group (n=176)		Δ
Demographic characteristics							
Gender			♀=61%		♀=61.9%		- ¹
Age (years)			μ =40.8 σ =18.7		μ =41.8 σ =18.7		- ²
(Net) income Indiv (EUR)			μ =1,321.9 σ =790.7		μ =1,428.9 σ =1,076.2		- ²
Shopping spending/income (%)			μ =40 σ =23		μ =41.8 σ =24.1		- ²
# of persons in hh			μ =2.4 σ =1.3		μ =2.2 σ =1		- ²
Education			A-level=33.8%		A-level=37.5%		- ¹
[Top 3]			Sec. school=21.3%		Sec. school=17.5%		
			University=20.6%		University=15.9%		
Profession			White collar worker=32.4%		White collar worker=27.8%		- ¹
[Top 3]			Student=23.5%		Senior citizen=27.8%		
			Senior citizen=19.9%		Student=19.9%		
Shopping behavior							
Shopping frequency per month in general			μ =15.4 σ =8.5		μ =13.5 σ =7		- ²
Visiting frequency per month			μ =6.8 σ =7.4		μ =6 σ =5.8		- ²
Expenditures (EUR) per visit			μ =40.5 σ =38.9		μ =39.1 σ =30		- ²
Retention time (min) per visit			μ =63.6 σ =43.3		μ =54.8 σ =36.4		- ²
Shops visited per trip			μ =2.8 σ =2.6		μ =2.6 σ =1.5		- ²
Caption: μ ...mean value; σ ...standard deviation; Δ ...significant difference; n...sample size; 1... χ^2 -Test, 2...Mann-Whitney-U-Test; -...no significant difference (p>.05); hh...household; indiv...individual; EUR...Euro; min...minutes; ♀...female;							

Table 3: Results from the 'One group pretest-posttest design'

Hypothesis	Independent variable (aroma)	Dependent variable	Δ^1
H1	Aroma has an impact on	the mall perception	partly accept/reject
H1a	The perception of ... is different after the visit [A] and before the visit [B].	retail tenant mix	** [E↑]
H1b		non-retail tenant mix	* [E↑]
H1c		price/value ratio	** [E↑]
H1d		product range	--
H1e		personnel	** [E↑]
H1f		smell	--
H1g		Other atmospherics	--
H2	Aroma has an impact on	customers' emotions	reject
H2a	The state of ... is different after the visit [A] and before the visit [B].	arousal	--
H2b		pleasure	--
H3	Aroma has an impact on	customers shopping behaviour	reject
H3a	The planned spending for goods	actual spending	-
H3b	The planned spending for food/drinks	is different to the actual spending	** [ACT↓]
H3c	The planned retention time	actual retention time	* [ACT↓]
H3d	The planned number of stores to be visited	actual of number of stores visited	* [ACT↓]
Caption: 1...Wilcoxon-Test, --... $p>.1$; -... $p<.1$; *... $p<.05$; **... $p<.01$; ***... $p<.001$; [E↑]...ratings/values of the experimental group are higher; [E↓]... ratings/values of the experimental group are lower; ACT↓...actual behaviour shows lower figures			

Table 4: Results from the 'Posttest only control group design'

Hypothesis	Independent variable (aroma)	Dependent variable	Δ^1
H1	Aroma has an impact on	the mall perception	partly accept/reject
H1a	The experimental group [E] differs from the control group [C] with respect to the perception of the	retail tenant mix	* [E↑]
H1b		non-retail tenant mix	* [E↑]
H1c		price/value ratio	--
H1d		product range	** [E↑]
H1e		personnel	*** [E↑]
H1f		smell	--
H1g		other atmospherics	* [E↑]
H2	Aroma has an impact on	customers' emotions	partly accept/reject
H2a	The experimental group [E] differs from the control group [C] with respect to their state of	arousal	** [E↑]
H2b		pleasure	--
H3	Aroma has an impact on	customers actual shopping behaviour	reject
H3a	The experimental group [E] differs from the control group [C] with respect to	actual spending for goods	--
H3b		actual spending for food/drinks	--
H3c		actual retention time	--
H3d		perceived retention time	--
H3e		actual number of stores visited	** (E↓)
H3e		actual number of stores visited where money was spent	--

Caption: 1...Mann-Whitney U-Test, --...p>.1; -...p<.1; *...p<.05; **...p<.01; ***...p<.001; [E↑]...ratings/values of the experimental group are higher; (E↓)... ratings/values of the experimental group are lower;

Table 5: Pretest-posttest control group design'

Hypothesis	Independent variable (aroma)	Dependent variable	Δ^1
H1	Aroma has an impact on	the mall perception	reject
H1a	The experimental groups differs from the control group with respect to changes of perception of the ...	retail tenant mix	--
H1b		non-retail tenant mix	-
H1c		price/value ratio	--
H1d		product range	--
H1e		personnel	--
H1f		odour	--
H1g		Other atmospherics	--
H2	Aroma has an impact on	customers' emotions	reject
H2a	The experimental groups differs from the control group with respect to changes of their state of ...	arousal	--
H2b		pleasure	--
H3	Aroma has an impact on	customers actual shopping behaviour	reject
H3a	The experimental groups differs from the control group with respect to changes between	Planned and actual spending	--
H3b		Planned and actual retention time	--
H3c		Planned and actual number of stores (to be) visited	* (E↓)

Caption: 1...Mann-Whitney U-Test, --...p>.1; -...p<.1; *...p<.05; **...p<.01; ***...p<.001; (E↓)... ratings/values of the experimental group are lower;

Table 6: Synopsis of hypotheses testing

Experimental design	One group pretest-posttest design	After-only with control group	Before-after with control group design
Hypotheses			
Aroma has an impact on ...			
H ₁ : perception of the mall	supported (partially) rejected	supported (partially)	rejected
H ₂ : state of emotions	rejected	supported (partially)	rejected
H ₃ : shopping behaviour	rejected	rejected	rejected

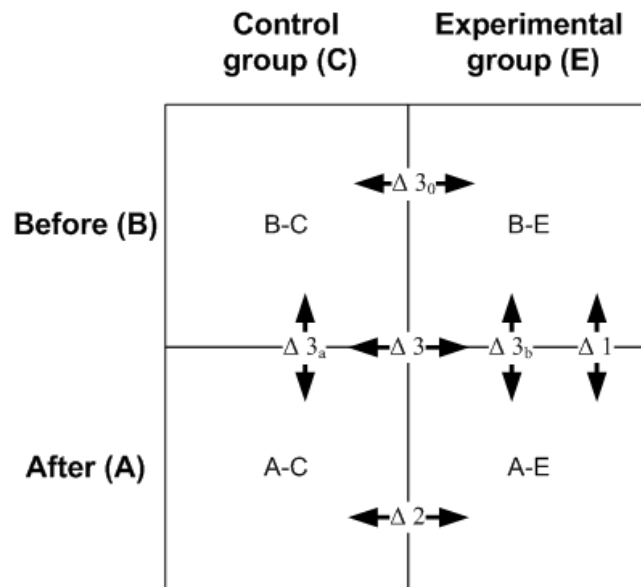


Figure 1: Comparison of the experimental designs used in this study

Appendix

Factor Indicator	Pre-measurement		Post-Measurement	
	Control Group	Experimental group	Control group	Experimental group
Retail tenant mix^a This mall has a broad range of retail stores. This mall has an attractive range of retail stores. Many well-known retail stores are in this mall.	$\alpha=.830$	$\alpha=.797$	$\alpha=.859$	$\alpha=.872$
Non-retail tenant mix^a This mall has a broad range of bars and restaurants. This mall offers a broad range of entertainment facilities.	$\alpha=.700$	$\alpha=.627$	$\alpha=.650$	$\alpha=.675$
Price-value ratio^a The overall price level is low in this mall. You can find a lot of special offers in this mall. The price-quality ratio is good in this mall.	$\alpha=.780$	$\alpha=.726$	$\alpha=.776$	$\alpha=.724$
Product range^a The quality of products offered in this mall is good. A broad range of products are offered in this mall. A large variety of products in each category is offered in this mall. A broad range of brands are available in this mall.	$\alpha=.844$	$\alpha=.768$	$\alpha=.884$	$\alpha=.805$
Personnel^a Personnel are friendly in this mall. Personnel are competent in this mall. Personnel are helpful in this mall.	$\alpha=.883$	$\alpha=.940$	$\alpha=.892$	$\alpha=.952$
Atmospheric stimuli^a The architecture of this mall is appealing. It smells pleasantly in this mall. The air is pleasant in this mall. The temperature is pleasant in this mall.	$\alpha=.819$	$\alpha=.843$	$\alpha=.840$	$\alpha=.864$

It is pleasantly bright in this mall.

The colour-design of this mall is likable.

This mall is always clean.

There is a good mood in this mall

There is a pleasant atmosphere in this mall.

Pleasure^b	$\alpha = .816$	$\alpha = .843$	$\alpha = .852$	$\alpha = .918$
Unhappy-happy				
Annoyed-pleasant				
Unsatisfied-satisfied				
Melancholic/contented				
Arousal^b	$\alpha = .634$	$\alpha = .642$	$\alpha = .610$	$\alpha = .683$
Relaxed-stimulated				
Calm-excited				

Caption: α ...Cronbach's Alpha, α ...seven point rating scale (anchors 0-6; totally disagree – totally agree); b ...seven point rating scale (anchors -3 to+3)